Approximate comparison of distance automata

Thomas Colcombet, Laure Daviaud

Approximate comparison of distance automata ⋆

Thomas Colcombet and Laure Daviaud

Université Sorbonne Paris Cité, Cnrs, Liafa

Abstract. Distance automata are automata weighted over the semiring $(\mathbb{N}\cup\{\infty\}, \min, +)$ (the tropical semiring). Such automata compute functions from words to $\mathbb{N}\cup\{\infty\}$ such as the number of occurrences of a given letter. It is known that testing $f \leqslant g$ is an undecidable problem for f, g computed by distance automata. The main contribution of this paper is to show that an approximation of this problem becomes decidable.
We present an algorithm which, given $\varepsilon > 0$ and two functions f, g computed by distance automata, answers "yes" if $f \leqslant (1-\varepsilon)g$, "no" if $f \not\leqslant g$, and may answer "yes" or "no" in all other cases. This result highly refines previously known decidability results of the same type.
The core argument behind this quasi-decision procedure is an algorithm which is able to provide an approximated finite presentation to the closure under product of set of matrices over the tropical semiring.
We also provide another theorem, of affine domination, which shows that previously known decision procedures for cost-automata have an improved precision when used over distance automata.

1 Introduction

One way to see language theory, and in particular the theory of regular languages, is as a toolbox of constructions and decision procedures allowing high level handling of languages. These high level operations can then be used as black-boxes in various decision procedures, such as in verification.

Since the early times of automata theory, the need for the effective handling of functions rather than sets (as languages) was already apparent. Schützenberger proposed already in the sixties models of finite state machines used for computing functions. These are now known as weighted automata [9] and are the subject of much attention from the research community. In general, weighted automata are non-deterministic automata, weighted over some semiring $(S, \oplus, \otimes)$. The value computed by such an automaton over a given word is then the sum (for $\oplus$) over every run over this word of the product (for $\otimes$) of the weights along the run.

Several instances of this model are very relevant for modelling the behaviour of systems, and henceforth attract much attention. This is in particular the case of probabilistic automata (over the semiring $(\mathbb{R}^+, +, \times)$ with some additional constraints enforcing weights to remain in $[0, 1]$), and distance automata which

⋆ The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n^{o}259454.

are automata weighted over the semiring $(\mathbb{N} \cup \{\infty\}, \min, +)$. In such an automaton, each transition is labelled with a non-negative integer (usually 0 or 1), and the weight of a word is the minimum over all possible paths of the sum of the weights. These automata naturally capture some optimisation problems since computing the value amounts to find the path of minimal weight.

The subject of this paper is to develop algorithmic tools for distance automata, and more precisely to develop the question of comparing distance automata. We know from the beginning that exact comparison is beyond reach.

Theorem 1 (Krob [6]). *The problem to determine, given two functions f, g computed by distance automata, whether $f = g$ or not is undecidable. The problem whether $f \leqslant g$ or not is also undecidable, even if g is deterministic.*

Despite this, some positive results exist but for a comparison relation less precise than inequality, namely the domination. Given two functions $A^* \to \mathbb{N} \cup \{\infty\}$, f *is dominated* by g (and we note $f \preccurlyeq g$) if there is a function $\alpha : \mathbb{N} \to \mathbb{N}$, extended with $\alpha(\infty) = \infty$, such that

$$f \leqslant \alpha \circ g \ .$$

Moreover, if α is a polynomial, we say that f *is polynomially dominated by* g. The following theorem shows the good properties of the domination relation.

Theorem 2 ([2] extending results and techniques from [4,8,11,5,1]). *Given two functions computed by distance automata, domination is decidable. Furthermore, if a function dominates another, then it polynomially dominates it*[1]*.*

The motivation of this work is to improve Theorem 2, and answer the following question:

> Is it possible to decide "approximations" of the inequality of functions computed by distance automata that are finer than domination ?

We answer positively this question in two ways. We first show:

Theorem 3 (affine domination). *Given two functions f and g computed by distance automata, if f is dominated by g then f is affinely dominated by g, i.e., $f \leqslant \alpha \circ g$ for some polynomial α of degree* 1.

A consequence of this theorem is that the decision procedure provided by Theorem 2 in fact decides the affine domination, which is finer than the polynomial domination[2].

Our second, and main contribution is an even more accurate decision-like procedure. One says that an algorithm, given two functions f and g and some real $\varepsilon > 0$, *ε-approximates the inequality* if:

---

[1] Technically, this is not stated in [2] , but can be derived directly from the proofs which explicitly compute the function α using operations preserving polynomials.

[2] Theorem 2 holds for more general classes of automata, cost automata, for which affine domination does not hold. Affine domination is specific to distance automata.

– if $f \leqslant (1-\varepsilon)g$, the output is "yes",
– if $f \not\leqslant g$, the output is "no",
– otherwise the output can be either "yes" or "no".

Hence, if such an algorithm answers "yes", one has a guaranty that $f \leqslant g$. Conversely if f is ε-inferior to g (meaning $f \leqslant (1-\varepsilon)g$), one is sure that the algorithm answers "yes". Our second and main result reads as follows:

Theorem 4 (approximate comparison). *There is an EXPSPACE algorithm which ε-approximates the inequality of functions computed by distance automata.*

This result is in fact a consequence of a theorem – called the core theorem below – stating that it is possible, given a set of matrices X in the tropical semiring, to approximate (in a suitable way) the set

$$\left\{\frac{1}{k}(M_1 \otimes \cdots \otimes M_k) \ : \ M_1, \ldots, M_k \in X\right\} ,$$

where $\otimes$ denotes the product of matrices. More precisely, the core theorem states that it is possible to approximate the upper envelope of the set of pairs $\{(M_1 \otimes \cdots \otimes M_k, k) \ : \ M_1, \ldots, M_k \in X\}$ for a suitable notion of approximation. This core theorem, Theorem 5, will be described precisely in the first section of this paper.

In Section 2 we present some classical definitions and formally state our core theorem. Section 3 is devoted to the proof of the core theorem. Section 4 applies the core theorem for answering our original motivation, and shows the decidability of the approximate comparison between distance automata. We prove on the way our result of affine domination, Theorem 3. Section 5 concludes the paper.

2 Description of the core theorem

In this first section, we introduce the basic definitions, and define sufficient material for stating our core theorem 5. Its proof is the subject of Section 3 and its application to the comparison of distance automata is the subject of Section 4.

We first introduce some classical algebraic definitions in Section 2.1, and finally state our core theorem in Section 2.2.

2.1 Classical definitions

A *semigroup* $(S, \cdot)$ is a set S equipped with an associative binary operation "$\cdot$". If the product has furthermore a neutral element, it is called a *monoid*. The monoid is said commutative when $\cdot$ is commutative. An *idempotent* in a monoid is an element e such that $e \cdot e = e$. Given a subset A of a semigroup, $\langle A \rangle$ denotes the closure of A under product, *i.e.,* the least sub-semigroup that contains A. Given two subsets X, Y of a semigroup, $X \cdot Y$ denotes the set $\{a \cdot b \ : \ a \in X,\ b \in Y\}$.

A *semiring* is a set S equipped with two binary operations $\oplus$ and $\otimes$ such that $(S, \oplus)$ is a commutative monoid of neutral element 0, $(S, \otimes)$ is a monoid of neutral element 1, 0 is absorbing for $\otimes$ (i.e., $x \otimes 0 = 0 \otimes x = 0$) and $\otimes$ distributes over $\oplus$. We will consider three semirings: $(\mathbb{R}^+ \cup \{\infty\}, \min, +)$, denoted $\overline{\mathbb{R}^+}$, its restriction to $\mathbb{N} \cup \{\infty\}$, denoted $\overline{\mathbb{N}}$, and its restriction to $\{0, \infty\}$ denoted $\mathbb{B}$. The third, finite semiring is called the *Boolean semiring*, since if we identify 0 with "true" and ∞ with "false", then $\oplus$ is the disjunction and $\otimes$ the conjunction. Remark that in the three cases, the "0" is ∞, and the "1" is 0.

Let S be one of the above semirings. The set of matrices with m rows and n columns over S is denoted $\mathcal{M}_{m,n}(S)$. For $M \in \mathcal{M}_{m,n}(S)$, we denote by $\widetilde{M}$ the matrix over $\mathbb{B}$ in which all entries of M different from ∞ are changed into 0. We define the multiplication $A \otimes B$ of two matrices A, B (provided the number n of columns of A equals the number of rows of B) as usual by:

$$(A \otimes B)_{i,j} = \bigoplus_{0<k\leqslant n} (A_{i,k} \otimes B_{k,j}) = \min_{0<k\leqslant n} (A_{i,k} + B_{k,j}) \ .$$

For a positive integer k, we also use the notation $M^k = \underbrace{M \otimes \cdots \otimes M}_{k \text{ times}}$.

For $\lambda \in S$, we denote by λA the matrix such that $(\lambda A)_{i,j} = \lambda A_{i,j}$, with the convention $\lambda\infty = \infty$ (the standard product is used here, not the one of the semiring). Finally, we denote by $B+\lambda$ the matrix such that $(B+\lambda)_{i,j} = B_{i,j} + \lambda$.

2.2 Weighted matrices and the core theorem

In this section we state our core approximation result, Theorem 5. This theorem states that given a set of weighted matrices, it is possible to compute a finite presentation of its closure under product up to some approximation. Hence we have to introduce weighted matrices, the approximation, and what are finite presentations before disclosing the statement. This requires some specific definitions that we present beforehand. In the following, a positive integer n is fixed, and all matrices implicitly belong to $\mathcal{M}_{n,n}(\overline{\mathbb{R}^+})$.

As already mentioned in the introduction, our goal is to approximate a set of pairs (M, ℓ) where M is a matrix and ℓ is a positive integer. We call such pairs weighted matrices. A *weighted matrix* is an ordered pair (M, ℓ) where $M \in \mathcal{M}_{n,n}(\overline{\mathbb{R}^+})$ and $\ell \in \mathbb{N}$ is non-null. The positive integer ℓ is called the *weight* of the weighted matrix. The set of weighted matrices is denoted by $\mathcal{W}_{n,n}$. Weighted matrices have a semigroup structure $(\mathcal{W}_{n,n}, \otimes)$, where $(M, \ell) \otimes (M', \ell')$ stands for $(M \otimes M', \ell + \ell')$. Given A, B subsets of $\mathcal{W}_{n,n}$, one denotes by $A \otimes B$ the set $\{M \otimes N \ : \ M \in A, \ N \in B\}$, and by $\langle A \rangle$ the closure under $\otimes$ of A. With this terminology, our goal is, given a finite set of weighted matrices X, to approximate $\langle X \rangle$.

We describe now the notion of approximation that we use. Given some $\varepsilon > 0$ and two weighted matrices (M, ℓ) and (M', ℓ'), one writes

$$(M, \ell) \preccurlyeq_\varepsilon (M', \ell') \qquad \text{if} \qquad \ell \geqslant \ell', \widetilde{M} = \widetilde{M'} \text{ and } M \leqslant M' + \varepsilon\ell \ .$$

Remark that in particular, this implies $\frac{1}{\ell}M \leqslant \frac{1}{\ell'}M' + \varepsilon$, which is the intention behind this definition. The definition of $\preccurlyeq_\varepsilon$ is more constraining: this is mandatory for having better properties with respect to the product of matrices, such as in Lemma 1 below. This definition extends to sets of weighted matrices as follows. Given two such sets X, X', $X \preccurlyeq_\varepsilon X'$ if for all $(M,\ell) \in X$, there exists $(M',\ell') \in X'$ such that $(M,\ell) \preccurlyeq_\varepsilon (M',\ell')$. One writes $X \approx_\varepsilon X'$ if $X \preccurlyeq_\varepsilon X'$ and $X' \preccurlyeq_\varepsilon X$ (and says X is ε-equivalent to X').

The following lemma establishes some simple properties of the $\preccurlyeq_\varepsilon$ relations (as a consequence, the same properties hold for $\approx_\varepsilon$).

Lemma 1. *Given $X, X', Y, Y', Z \subseteq \mathcal{W}_{n,n}$ and $\varepsilon, \eta > 0$,*

- *if $X \preccurlyeq_\varepsilon Y$ and $Y \preccurlyeq_\eta Z$ then $X \preccurlyeq_{\varepsilon+\eta} Z$,*
- *if $X \preccurlyeq_\varepsilon X'$ and $Y \preccurlyeq_\varepsilon Y'$ then $X \otimes Y \preccurlyeq_\varepsilon X' \otimes Y'$,*
- *if $X \preccurlyeq_\varepsilon X'$ then $\langle X \rangle \preccurlyeq_\varepsilon \langle X' \rangle$.*

Proof. First item. If $(M,\ell) \preccurlyeq_\varepsilon (M',\ell') \preccurlyeq_\eta (M'',\ell'')$, then $\ell \geqslant \ell' \geqslant \ell''$, $\tilde{M} = \tilde{M}' = \tilde{M}''$ and $M \leqslant M' + \varepsilon\ell \leqslant M'' + \eta\ell' + \varepsilon\ell \leqslant M'' + (\varepsilon+\eta)\ell$. This easily extends to sets of weighted matrices.

Second item. Assume $(M,\ell) \preccurlyeq_\varepsilon (M',\ell')$ and $(N,t) \preccurlyeq_\varepsilon (N',t')$. Then, $\ell+\ell' \geqslant t+t'$, $\widetilde{M \otimes N} = \widetilde{M' \otimes N'}$ and $M \otimes N \leqslant (M'+\varepsilon\ell) \otimes (N'+\varepsilon t) \leqslant M' \otimes N' + \varepsilon(\ell+t)$. This naturally extends to sets of weighted matrices.

Third item. By induction, applying the second item. ☐

The last ingredient required is to describe how to represent (infinite) sets of weighted matrices. Call a set of weighted matrices $W \subseteq \mathcal{W}_{n,n}$ *finitely presented* if it is a finite union of singleton sets, and of sets of the form $\{(kM,k) : k \geqslant \ell\}$ where $M \in \mathcal{M}_{n,n}(\overline{\mathbb{R}^+})$ and ℓ is a positive integer. Our algorithm manipulates finitely presented sets of weighted matrices.

The core technical contribution of this paper can now be stated, as follows.

Theorem 5 (core theorem). *Given $X \subseteq \mathcal{W}_{n,n}$ finitely presented and $\varepsilon > 0$, one can compute effectively $Y \subseteq \mathcal{W}_{n,n}$ finitely presented such that:*

$$Y \approx_\varepsilon \langle X \rangle \ .$$

A sketch of the proof of this result will be the subject of Section 3. The application of this theorem to the comparison of distance automata is presented in Section 4. The two sections are independent.

3 Proof of the core theorem

In this section we describe the key arguments involved in the proof of Theorem 5. It is the combination of several arguments. The first one is the use of the factorisation forest theorem of Simon.

3.1 The main induction: the forest factorization theorem of Simon

The forest factorization theorem of Simon [10] is a powerful combinatorial tool for understanding the structure of finite semigroups. In this short abstract, we will not describe the original statement of this theorem, in terms of trees of factorisations, but rather a direct consequence of it which is central in our proof.

Proposition 1 (equivalent to the forest factorization theorem [10][3]). *Given a semigroup morphism ϕ from $(S, \otimes)$ (possibly infinite) to a finite semigroup $(T, \cdot)$, and some $X \subseteq S$, set $X_0 = X$ and for all $k \geqslant 0$,*

$$X_{k+1} = X_k \cup X_k \otimes X_k \cup \bigcup_{e \cdot e = e \in T} \langle X_k \cap \phi^{-1}(e) \rangle \ ,$$

then $X_N = \langle X \rangle$ for $N = 3|T| - 1$.

This proposition teaches us that, for computing the closure under product in the semigroup S, it is sufficient to know how to compute (a) the union of sets, (b) the product of sets, and (c) the restriction of a set to the inverse image of an idempotent by ϕ, and (d) the closure under product of sets of elements that all have the same idempotent image under ϕ. Of course, this proposition is interesting when the semigroup T is cleverly chosen.

In our case, we are going to use the above proposition with $(S, \otimes) = \mathcal{W}_{n,n}$, $(T, \cdot) = \mathcal{M}_{n,n}(\mathbb{B})$, and ϕ the morphism which maps (M, ℓ) to $\widetilde{M}$. Our algorithm will compute, given a finitely presented set of weighted matrices X, an approximation of $\langle X \rangle$ following the same inductive construction as in the forest factorisation theorem. This is justified by the two following lemmas, that we prove below.

Lemma 2. *For all $\varepsilon > 0$ and all finitely presented $X, Y \subseteq \mathcal{W}_{n,n}$ there exists effectively* $\texttt{product}(\varepsilon, X, Y) \subseteq \mathcal{W}_{n,n}$ *finitely presented such that*

$$\texttt{product}(\varepsilon, X, Y) \approx_\varepsilon X \otimes Y \ .$$

Let X be a set of weighted matrices, we denote $\widetilde{X} = \{\widetilde{M} \mid \exists \ell > 0 \text{ s.t } (M, \ell) \in X\}$.

Lemma 3. *For all $\varepsilon > 0$ and all finitely presented $X \subseteq \mathcal{W}_{n,n}$ such that $\widetilde{X} = \{e\}$ for an idempotent e, there exists effectively* $\texttt{idempotent}(\varepsilon, X) \subseteq \mathcal{W}_{n,n}$ *finitely presented such that*

$$\texttt{idempotent}(\varepsilon, X) \approx_\varepsilon \langle X \rangle \ .$$

Assuming that Lemmas 2 and 3 hold, it is easy to provide an algorithm which, given $X \subseteq \mathcal{W}_{n,n}$ finitely presented, computes $X' \subseteq \mathcal{W}_{n,n}$ finitely presented such that $X' \approx_\varepsilon \langle X \rangle$. The principle of the algorithm is to implement Proposition 1, using finitely presented sets that approximate the X_k's.

[3] Modern proofs of this theorem can be found in [7,3], in particular with the exact bound of $N = 3|T| - 1$ (Simon's original proof only provides $N = 9|T|$).

- Set $Y_0 = X$ and $N = 3(2^{n^2}) - 1$.
- For all $0 \leqslant k \leqslant N$, set $\varepsilon(k) = \frac{\varepsilon}{2^{N-k}}$ and

$$Y_{k+1} = Y_k \cup \mathtt{product}(\varepsilon(k), Y_k, Y_k) \cup \bigcup_{e \otimes e = e \in T} \mathtt{idempotent}(\varepsilon(k), Y_k \cap \phi^{-1}(e)) \ .$$

- output Y_N

It is easy to prove that this construction is correct. Indeed, one proves by induction that $Y_k \approx_{\varepsilon(k)} X_k$ for all $k = 0, \ldots, N$ where X_k is defined as in Proposition 1 (with $S = \mathcal{W}_{n,n}$, $T = \mathcal{M}_{n,n}(\mathbb{B})$ and $\phi(M, \ell) = \widetilde{M}$). For $k = 0$, one has $X_k = X = Y_k$. Let $k \geqslant 0$, suppose that $Y_k \approx_{\varepsilon(k)} X_k$, then by Lemma 2, Lemma 1 and the induction hypothesis,

$$\mathtt{product}(\varepsilon(k), Y_k, Y_k) \approx_{\varepsilon(k)} Y_k \otimes Y_k \approx_{\varepsilon(k)} X_k \otimes X_k \ .$$

Finally, by Lemma 1, $\mathtt{product}(\varepsilon(k), Y_k, Y_k) \approx_{2\varepsilon(k)} X_k \otimes X_k$. Similarly, by Lemma 3, for all idempotent e, $\mathtt{idempotent}(\varepsilon(k), Y_k \cap \phi^{-1}(e)) \approx_{2\varepsilon(k)} \langle X_k \cap \phi^{-1}(e)\rangle$. Thus $Y_{k+1} \approx_{\varepsilon(k+1)} X_{k+1}$.

Hence, what remains to be done is to establish Lemmas 2 and 3.

3.2 Approximate products of sets

The proof of Lemma 2 shows explicit examples of the approximation arguments that are later used in a more advanced way.

Proof (Proof of Lemma 2). Since the finitely presented sets of weighted matrices are closed under union, it is sufficient to prove Lemma 2 for the atomic blocks of the finite presentation. Namely, it is sufficient to consider the case $X = \{(M, x)\}$ or $X = \{(\ell M, \ell) \mid \ell \geqslant x\}$ together with $Y = \{(N, y)\}$ or $Y = \{(\ell N, \ell) \mid \ell \geqslant y\}$. This results in four possibilities, among which only three remain up to symetry: (a) $X = \{(M, x)\}$ and $Y = \{(N, y)\}$, (b) $X = \{(M, x)\}$ and $Y = \{(\ell N, \ell) \mid \ell \geqslant y\}$, and finally (c) $X = \{(\ell M, \ell) \mid \ell \geqslant x\}$ and $Y = \{(\ell N, \ell) \mid \ell \geqslant y\}$.

For space reason, let us just explain the most interesting case, case (c). Let a be the maximum absolute value of a non-infinite entry of M or N. Choose some z such that $2ax \leqslant \varepsilon z$ and $2ay \leqslant \varepsilon z$, and let Z be the set $Z_1 \cup Z_2$ defined by:

$$\begin{aligned} & Z_1 = \{(x' M \otimes y' N, x' + y') \mid x' + y' < z\} \ , \\ \text{and} \qquad & Z_2 = \{(\ell(\lambda M \otimes (1 - \lambda) N), \ell) \mid \ell \geqslant z,\ \lambda \in [0, 1]\} \ . \end{aligned}$$

The set Z_1 is finite, and merely lists all weighted matrices of weight less than z in $X \otimes Y$. The set Z_2 (which is not finitely presented) takes all barycentres of M and N, and produces corresponding weighted matrices for all possible weights greater or equal to z. We need to prove two things. First that $Z \approx_{\frac{\varepsilon}{2}} X \otimes Y$, and second that one can further approximate Z_2 by a finitely presented $Z_3 \approx_{\frac{\varepsilon}{2}} Z_2$. By Lemma 1 we can then conclude that $X \otimes Y \approx_\varepsilon Z_1 \cup Z_3$, and that $Z_1 \cup Z_3$ is finitely presented and computable from X and Y.

Let us prove that $Z \approx_{\frac{\varepsilon}{2}} X \otimes Y$. Remark first that $X \otimes Y \subseteq Z$. For the converse direction, consider $(W, \ell) \in Z$. Clearly, if $\ell < z$, then $(W, \ell) \in Z_1 \subseteq X \otimes Y$. Otherwise, $W = (\lambda \ell M) \otimes ((1-\lambda)\ell N))$. It is sufficient for us to find $x' \geqslant x$ and $y' \geqslant y$ such that $x' + y' = \ell$, and $\left|\lambda - \frac{x'}{\ell}\right| \leqslant \frac{\varepsilon}{2a}$: indeed, assuming the existence of such x', y', the matrix $W' = (x'M \otimes y'N, \ell)$ is such that $(W, \ell) \approx_{\frac{\varepsilon}{2}} (W', \ell)$, and furthermore $(W', \ell) \in X \otimes Y$. For proving the existence of such x', y', consider the evolution of the value $\frac{x'}{\ell}$ when x' ranges from x to $\ell - y$. Since $\ell \geqslant z$, $\frac{x}{\ell} \leqslant \frac{\varepsilon}{2a}$, and similarly $\frac{\ell - y}{\ell} \geqslant 1 - \frac{\varepsilon}{2a}$. Furthermore when x' increases of 1, the quantity $\frac{x'}{\ell}$ increases of at most $\frac{1}{z} \leqslant \frac{\varepsilon}{2a}$. As a consequence, $\frac{x'}{\ell}$ gets to be $\frac{\varepsilon}{2a}$-close of any $\lambda \in [0, 1]$ when x' ranges from x to $\ell - y$. Consider x' witnessing this fact and set $y' = \ell - x'$. The pair x', y' satisfies the requirement.

One now needs defining a set $Z_3 \approx_{\frac{\varepsilon}{2}} Z_2$ which is finitely presented. The set Z_3 is defined as the set Z_2, but for the fact that λ is discretized by steps of $\frac{\varepsilon}{4a}$. This can be written as:

$$Z_3 = \bigcup_{\lambda \in \left([0,1] \cap \frac{\varepsilon}{4a}\mathbb{N}\right)} \{(\ell(\lambda M \otimes (1-\lambda)N), \ell) \mid \ell \geqslant z\} \ .$$

Clearly, this set is finitely presented. It is also simple to prove that $Z_3 \approx_{\frac{\varepsilon}{2}} Z_2$. $\square$

Details of the complete proof are given in Section A.

3.3 Approximate "idempotent" products of sets

We enter here the most technical part of the proof. We just sketch here the essential ideas. Let us fix ourselves an idempotent $E \in \mathcal{M}_{n,n}(\mathbb{B})$, some $\varepsilon > 0$, and some finitely presented set of weighted matrices X such that $\widetilde{X} = \{E\}$. Our goal is to construct a finitely presented X' such that $X' \approx_\varepsilon \langle X \rangle$. In this section, all matrices and weighted matrices M are supposed to be such that $\widetilde{M} = E$.

The proof is done in several steps. We first define two restricted notions of closure. Let p be some positive integer and $\eta > 0$. Define $\langle X \rangle_{p,\eta}$ to be the set of weighted matrices

$$(M, \ell) = (M_1, \ell_1) \otimes \cdots \otimes (M_k, \ell_k)$$

where each (M_i, ℓ_i) belongs to X, and there exists $i_1 < \cdots < i_s$ with $s \leqslant p$ such that $\sum_{j=1}^{s} \ell_{i_j} \geqslant (1-\eta)\ell$ (where $\ell = \sum_{j=1}^{k} \ell_j$). In other words, k can be very large, but there are only few weighted matrices (less than p) that count for most of the weight (ratio $1 - \eta$). If furthermore $\ell_1 \leqslant \eta\ell$ and $\ell_k \leqslant \eta\ell$, the product is said *uniform*. Set $\langle X \rangle^u_{p,\eta}$ to be the set of weighted matrices obtained by such uniform products.

The following lemma shows that $\langle X \rangle_{p,\eta}$ and $\langle X \rangle^u_{p,\eta}$ can be effectively approximated (for sufficiently small choices of η).

Lemma 4. *For all $\gamma > 0$, there exists $\eta > 0$ such that for all finitely presented X with $\widetilde{X} = \{E\}$ and all p, there exist effectively Y and Z finitely presented such that:*

$$Y \approx_\gamma \langle X \rangle_{p,\eta} \text{ and } Z \approx_\gamma \langle X \rangle^u_{p,\eta} \ .$$

The idea behind this proof is that the set $\langle X \rangle_{p,\eta}$ is essentially like a product of up to p weighted matrices (something we know how to do as shown in Lemma 2), but modified in order to possibly interleave in the product matrices of "small weights". One chooses η sufficiently small such that all these matrices of small weight can be approximated by matrices of "null weight".

Our second intermediate lemma shows that any product of weighted matrices in X can be decomposed into products as the above restricted closures.

Lemma 5. *For all X such that $\widetilde{X} = \{E\}$ for some idempotent E, for all $\eta > 0$, there is an integer p, such that:*

$$\langle X \rangle = \langle X \rangle_{p,\eta} \otimes \langle \langle X \rangle^u_{p,\eta} \rangle \otimes \langle X \rangle_{p,\eta} \ .$$

Here, the lemma comes from a careful analysis of the proportions between weights in any product $(M_1, \ell_1) \otimes \cdots \otimes (M_k, \ell_k)$.

At this point one should understand why computing $\langle \langle X \rangle^u_{p,\eta} \rangle$ is simpler than computing $\langle X \rangle$ as in the general case. The reason is that the weighted matrices in $\langle X \rangle^u_{p,\eta}$ have particularly good properties. Call a matrix M (such that $\widetilde{M} = E$) *uniform* if $M = E \otimes M \otimes E$. A weighted matrix (M, ℓ) is *uniform* if M is uniform. The following lemma shows that every matrix in $\langle X \rangle^u_{p,\eta}$ is "almost" uniform. Lemma 6 states it and refines Lemma 4.

Lemma 6. *For all $\gamma > 0$, there exists $\eta > 0$ such that for all finitely presented X with $\widetilde{X} = \{E\}$ and all p, there exist effectively Z finitely presented such that:*

$$Z \approx_\gamma \langle X \rangle^u_{p,\eta} \ ,$$

and all weighted matrices in Z are uniform.

Intuitively, this comes from the fact that weighted matrices in $\langle X \rangle^u_{p,\eta}$ are defined as products that starts and ends with matrices of "small weight". If these weights are sufficiently small, these matrices count as if of null weight. Thus, these extremity matrices act as multiplying on the left and on the right with E. This yields matrices that are very close to uniform ones.

Finally, one shows that it is possible, given a finitely presented set Y of uniform weight matrices, to compute its closure.

Lemma 7. *For all η and all finitely presented set Y of uniform weight matrices (with $\widetilde{Y} = E$), there exists effectively Z finitely presented such that:*

$$Z \approx_\eta \langle Y \rangle \ .$$

This results comes from the fact that one understands precisely the structure of uniform matrices (for instance, if we see such a matrix as a weighted graph, then the weight of transitions are constant in strongly connected component). Using this knowledge, it is possible to show that the "worst situation" occurs for very simple patterns of repetition of the matrices in Y. Since the relation $\approx_\eta$ just refers to the upper envelope, this worst situation is sufficient for us to conclude.

The combination of all the above lemmas yields quite simply a proof of Lemma 3. Details of the proof are given in Section B.

4 Comparing distance automata

In this section, we consider the problem of comparing the functions computed by distance automata. In particular, we establish Theorem 3, and we reduce Theorem 4 to our core theorem, Theorem 5.

We start by describing distance automata, and their relationship with matrices over the tropical semiring (Section 4.1).

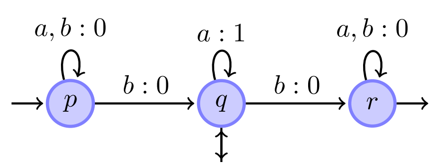
4.1 Distance automata

An *alphabet* is a finite set. The set of *words* over an alphabet $\mathbb{A}$ is denoted $\mathbb{A}^*$. The *empty word* is λ. A *distance automaton* is a tuple $(\mathbb{A}, Q, I, F, T)$, where Q is a finite set of *states* (that we can assume to be $\{1, \ldots, n\}$) where I (resp. T) is a row-vector (*resp.* column-vector) indexed by Q, and F is a morphism from words to $\mathcal{M}_{n,n}(\overline{\mathbb{N}})$. The function f computed by a distance automaton $(\mathbb{A}, Q, I, F, T)$ over an input word u is:

$$\begin{aligned} f : \mathbb{A}^* &\to \overline{\mathbb{N}} \\ u &\mapsto I \otimes F(u) \otimes T \ . \end{aligned}$$

We assume from now on that the initial and final vectors I, T of distance automata only range over $\{0, \infty\}$. The theorems are equally true without this assumption, but this simplifies slightly the proof. In practice the theorems without this restriction can be obtained by simple reductions to this case.

We have defined so far distance automata in terms of matrices. One can see this object in a more "automaton" form as follows. There is a transition labelled (a, x) from state p to state q if $x < \infty$ and $x = F(a)_{p,q}$. A state p is *initial* if $I_{1,p} = 0$. It is *final* if $T_{i,1} = 0$. An example of distance automaton is as follows:



One can redefine the function computed by a distance automaton as follows. A *run* of an automaton over a word $a_1 \ldots a_k$ is a sequence $p_0, \ldots, p_k$ of states. The *weight of a run* is the sum of the weights of its transitions, *i.e.*, $F(a_1)_{p_0,p_1} + \cdots + F(a_k)_{p_{k-1},p_k}$. Remark that if there is some non-existing transition in this sequence, say from p_{i-1} to p_i, this means that $F(a_i)_{p_{i-1},p_i} = \infty$, and as a consequence the run has an infinite weight. A run is *accepting* if p_0 is initial and p_k is final. One defines the function accepted by the automaton as:

$$\begin{aligned} f : \mathbb{A}^* &\to \overline{\mathbb{N}} \\ u &\mapsto \inf\{\text{weight}(\rho) \ : \ \rho \text{ accepting run over } u\} \ . \end{aligned}$$

This definition is equivalent to the matrix version presented above.

For instance, the function computed by the above automaton associates to each word $u = a^{n_0} b a^{n_1} \ldots b a^{n_k}$ the value $\min(n_0, \ldots, n_k)$.

4.2 Superior limits

In this section, we present Theorem 6. This result, that is a refinement of known proofs concerning distance automata, will prove useful for further reductions.

In order to define the superior limit of a set of matrices, a topology is required. The matrices over $\overline{\mathbb{N}}$ are equipped with the following topology. When two matrices are distinct, their distance is $1/n$ where n is maximal positive integer such that the entries that carry values at most n are the same in both matrices. If no such integer exists, the distance is 1.

Given $X \subseteq \mathcal{M}_{n,n}(\overline{\mathbb{N}})$, a matrix N belongs to the *superior limit* of X if:

- N is the limit of some sequence of matrices from X,
- there exists no $M \in X$ such that $M > N$.

Let us call $\limsup(X)$ the set of matrices in the superior limit of S.

Theorem 6 (consequence of [4,8]). *Given a set $X \subseteq \mathcal{M}_{n,n}(\overline{\mathbb{N}})$, $\limsup(X)$ is finite. Furthermore, there is a PSPACE algorithm which, given a morphism F from $\mathbb{A}^*$ to $\mathcal{M}_{n,n}(\overline{\mathbb{N}})$, and a language $L \subseteq \mathbb{A}^*$ enumerates* $\limsup(F(L))$.

The first part of the statement is a consequence of Higman's lemma. The second part is an adaptation of Leung's proof of decidability of limitedness for distance automata [8] (it subsumes this result). We are not aware of any similar statement in the literature, though it can be deduced from previous works.

4.3 A first reduction: the theorem of affine domination

Our goal in this section is to establish the theorem of affine domination (Theorem 3). This will be the opportunity to introduce some notations used in the subsequent section.

Let us fix ourselves two distance automata over the same alphabet $\mathbb{A}$. The first one, $\mathcal{A}_f = (\mathbb{A}, Q_f, F, I_f, T_f)$ calculates a function f. The second one, $\mathcal{A}_g = (\mathbb{A}, Q_g, G, I_g, T_g)$ calculates a function g.

Define $R_{p,0,q} \subseteq \mathbb{A}^*$ to be the set of words over which there is a run of g of weight 0 from state p to state q. Let ℓ be a non-null weight occurring in some transition of g, and p, q be states in Q_g. Define $R_{p,\ell,q} \subseteq \mathbb{A}^*$ to contain the words over which there is a run of g from state p to state q which uses one transition of weight ℓ, and otherwise only transitions of weight 0. We will reuse this languages in the next section.

Proof (Proof of theorem 3). Define K to be the largest number that occur in one of $\limsup(F(R_{p,\ell,q}))$ for some states p, q and weight of a transition ℓ (such a number exists since by Theorem 6 it is the maximum of finitely many numbers). Given a matrix M, call an *m-expansion* of M a matrix $M' \geqslant M$ such that for all i, j such that $M_{i,j} > K$, $M'_{i,j} \geqslant m$. We first show a claim concerning expansions.

Claim. For all $M \in F(R_{p,\ell,q})$ and all m there exists an m-expansion $M' \in F(R_{p,\ell,q})$ for M.
Indeed, by definition of the superior limit, there is some $L \in \limsup(F(R_{p,\ell,q}))$ such that $L \geqslant M$. Furthermore, by choice of K, whenever $M_{i,j} > K$, $L_{i,j} = \infty$. Finally, still by definition of the superior limit, L is the limit of a sequence of matrices in $F(R_{p,\ell,q})$. Hence, for all m, there exists a matrix M' in this sequence which is sufficiently close to L that it is an m-expansion of M. This proves the claim.

Let us turn now to the core of the proof. Our goal is to prove that if f is dominated by g, (*i.e.,* there exists $\alpha : \mathbb{N} \to \mathbb{N}$ extended with $\alpha(\infty) = \infty$ such that $f \leqslant \alpha \circ g$), then $f \leqslant K(1+g)$. The proof is by contraposition. Thus, assume $f \not\leqslant K(1+g)$. This means $f(u) > Kg(u)+K$ for some word u. We have to prove that f is not dominated by g.

The first case is $g(u) = 0$. This means that $u \in R_{p,0,q}$ with p initial and q final. Using the above claim, one can chose for all m a word $v^m \in R_{p,0,q}$ such that $F(v^m)$ is an m-expansion of $F(u)$. Since $f(u) > K$, this means that for all initial state r and all final state s of $\mathcal{A}_f$, $F(u)_{r,s} > K$. This means that for all such r, s, $F(v^m)_{r,s} \geqslant m$. It follows that $f(v^m) \geqslant m$. Hence over the sequence $(v^m)_m$, g is bounded and f tends to infinity. This forbids the existence of a function α such that $f \leqslant \alpha \circ g$, f is not dominated by g.

Assuming $g(u) \neq 0$, the argument is similar. Remark first that $g(u)$ is finite since $f(u) > Kg(u) + K$. This means one can find $p_0, \ldots, p_k$ with p_0 initial, p_k final, and such that:

$$u = u_1 \ldots u_k, \quad u_1 \in R_{p_0,\ell_1,p_1}, \ldots, u_k \in R_{p_{k-1},\ell_k,p_k} \ ,$$

where $\ell_1, \ldots, \ell_k$ are all non-null and of sum $g(u)$. By the above claim, for all $i = 1 \ldots k$, and all m, one can select v_i^m in R_{p_{i-1},ℓ_i,p_i} such that $F(v_i^m)$ is an m-expansion of $F(u_i)$. Consider now the word $v^m = v_1^m \ldots v_k^m$. Clearly $g(v^m) = g(u)$. For the sake of contradiction, assume now that $F(v^m) < m$ for some

m. This means that there exists $q_0, \ldots, q_k$ such that q_0 is initial, q_k is final, and $F(v_i^m)_{q_{i-1},q_i} < m$ for all $i = 1 \ldots k$. Since $F(v_i^m)$ is an m -expansion of $F(u_i)$, this implies $F(u_i)_{q_{i-1},q_i} \leqslant K$. It follows that $F(u) \leqslant Kk \leqslant Kg(u)$. A contradiction. Hence $f(v^m) \geqslant m$. Thus, g is bounded over $(v^m)_m$ while f is not. As a consequence, f is not dominated by g. □

4.4 The reduction construction

We reuse definitions and notations of automata $\mathcal{A}_f$ and $\mathcal{A}_g$ given in the preceding section. In particular, we use the sets $R_{p,\ell,q}$ again.

Our goal is to construct a finite set of weighted matrices X that captures the relationship between f and g. The key ideas behind this reduction are the following. Each matrix (M, ℓ) in X corresponds to a set of runs of g, that start in a given state p and end in a given state q, and use exactly one transition of non-null weight, of weight ℓ. The corresponding matrix M is in charge of (a) simulating the behaviour of F over some word corresponding to such a run (there may be infinitely many such runs, but only the finitely many matrices of the superior limit need be considered), and (b) keeping information concerning the first and last state of the run of $\mathcal{A}_g$ for being able to check that pieces of run of g are correctly concatenated.

One also needs to define the part of the matrix in charge of controlling the validity of the run of $\mathcal{A}_g$. The construction behind Lemma **??** below is the one of a deterministic automaton, that reads words over the alphabet Q_g^2, and accepts a word $(p_1, q_1) \ldots (p_k, q_k)$ if, either p_1 is not initial, or q_k is not final, or if $q_{i-1} \neq p_i$ for some i. One can verify that this language is accepted by a deterministic and complete automaton of states $Q_g \uplus \{i, \bot\}$. The unique initial state is i, and, when reading the word $(p_1, q_1) \ldots (p_k, q_k)$, the automaton reaches state $\bot$ if p_1 is not initial or $q_{i-1} \neq p_i$ for some i, otherwise it reaches state q_k. The final states are the one not in T_g plus $\bot$ plus possibly i if there are no states that are both initial and final in g. Translated in matrix form, this yield Lemma 8.

Lemma 8. *There are $(n+2, n+2)$-matrices $(C^{p,q})_{p,q \in G}$ over $\mathbb{B}$ and vectors I_g and T_g such that for all $p_1, q_1, \ldots, p_k, q_k \in Q_g$,*

$$I_C \otimes C^{p_1,q_1} \otimes \cdots \otimes C^{p_k,q_k} \otimes T_C = \begin{cases} \infty & \text{if } p_1 \in I_g,\ q_1 = p_2,\ \ldots,\ q_{k-1} = p_k \text{ and } q_k \in T_g, \\ 0 & \text{otherwise.} \end{cases}$$

We can now construct the set X as follows:

$$X = \left\{ \left(\begin{pmatrix} M & \infty \\ \infty & C^{p,q} \end{pmatrix}, \ell \right) \ :\ M \in \limsup(F(R_{p,\ell,q})) \right\}$$

and the vectors

$$I = (I_f\ I_C) \qquad \text{and} \qquad T = \begin{pmatrix} T_f \\ T_c \end{pmatrix} .$$

The following lemma shows the validity of the construction, and more particularly how it relates the comparison of distance automata to the computation of the closure of a set of weighted matrices.

Lemma 9. *For all $\beta > 0$, $f \leqslant \beta g$ if and only if for all $(W, \ell) \in \langle X \rangle$, $I \otimes W \otimes T \leqslant \beta\ell$.*

Proof. Assume first $f \nleqslant \beta g$, which means $f(u) > \beta g(u)$ for some u. Then clearly, $g(u)$ is finite and hence, there is an accepting run ρ of g over u. This means that one can find $p_0, \ldots, p_k$ with p_0 initial, p_k final, and such that:

$$u \in R_{p_0,\ell_1,p_1} R_{p_1,\ell_2,p_2} \ldots R_{p_{k-1},\ell_k,p_k} \ ,$$

where $\ell_1, \ldots, \ell_k$ are all non-null and of sum $\ell = g(u)$. For all $i = 1 \ldots k$, set M_i to be some matrix in $\limsup(F(R_{p_{i-1},\ell_i,p_i}))$ such that $F(u_i) \leqslant M_i$. Let also C_i be C^{p_{i-1},p_i}. Clearly, the weighted matrix

$$(W_i, \ell_i) \qquad\qquad \text{with} \quad W_i = \begin{pmatrix} M_i & \infty \\ \infty & C_i \end{pmatrix}$$

belongs to X. Hence (W, ℓ) belongs to $\langle X \rangle$, where $W = W_1 \otimes \cdots \otimes W_k$. We then have $I \otimes W \otimes T = \min(x_f, x_C)$ with

$$x_f = I_f \otimes M_1 \otimes \cdots \otimes M_k \otimes T_f \qquad \text{and} \quad x_C = I_C \otimes C_1 \otimes \cdots \otimes C_k \otimes T_C \ .$$

By choice of the M_i's, $x_f \geqslant I_f \otimes F(u) \otimes T_f = f(u)$. Furthermore, by Lemma 8, $x_C = \infty$. It follows that $I \otimes W \otimes T \geqslant f(u) > \beta g(u) = \beta\ell$.

Assume now that $f \leqslant \beta g$. Consider some $(W, \ell) \in \langle X \rangle$, it is obtained as $(W, \ell) = (W_1, \ell_1) \otimes \cdots \otimes (W_k, \ell_k)$ with $(W_i, \ell_i) \in X$ for all i. By definition of X, each of the W_i's can be written, for some $p_i, q_i \in Q_g$, as

$$W_i = \begin{pmatrix} M_i & \infty \\ \infty & C^{p_i,q_i} \end{pmatrix} \qquad\qquad \text{with} \quad M_i \in \limsup F(R_{p_i,\ell_i,q_i}).$$

Once more, one has $I \otimes W \otimes T = \min(x_f, x_C)$ with

$$x_f = I_f \otimes M_1 \otimes \cdots \otimes M_k \otimes T_f \qquad \text{and} \quad x_C = I_C \otimes C_1 \otimes \cdots \otimes C_k \otimes T_C \ .$$

Remark first that if $x_C = 0$, clearly, $I \otimes W \otimes T = 0 \leqslant \beta\ell$. Hence, let us assume that $x_C = \infty$. This means by Lemma 8 that either p_1 is initial, q_k is final, and $p_i = q_{i-1}$ for all $i = 2 \ldots k$. One needs to prove $x_f \leqslant \beta\ell$.

Assume for the sake of contradiction that $x_f > \beta\ell$. By continuity of the product, and using the definition of the superior limit, there exist words $u_1, \ldots, u_k$ such that for all $i = 1 \ldots k$, $u_i \in R_{p_i,\ell_i,q_i}$, and $I_f \otimes F(u_1) \otimes \cdots \otimes F(u_k) \otimes T_f > \beta\ell$. Furthermore, by definition of the sets R_{p_i,ℓ_i,q_i}, the fact that p_1 is initial, that q_k is final, and that $q_{i-1} = p_i$ for all $i = 2 \ldots k$, it follows that $g(u_1 \ldots u_k) = \ell$. It follows that $f(u_1 \ldots u_k) > \beta g(u_1 \ldots u_k)$. A contradiction. □

We are now ready to establish the main theorem of the paper.

Proof (Proof of Theorem 4). Let us consider two functions f and g computed by distance automata and some $\varepsilon > 0$. The algorithm works as follows. It computes the set X of weighted matrices as defined in this section, as well as the corresponding vectors I, T. Using Theorem 5, it computes a finitely presented set Y of weighted matrices such that $Y \approx_{\frac{\varepsilon}{2}} \langle X \rangle$. Then it tests the existence in Y of a weighted matrix (M, ℓ) such that $I \otimes \frac{1}{\ell} M \otimes T > 1 - \frac{\varepsilon}{2}$. This is easy to do for finitely presented sets. If such a weighted matrix exists, the algorithm answers "no". It answers "yes" otherwise. Let us show the correctness of this approach.

- Assume $f \leqslant (1-\varepsilon)g$, and that, for the sake of contradiction, the algorithm answers "no". This means that $I \otimes \frac{1}{\ell} M \otimes T > 1 - \frac{\varepsilon}{2}$ for some weighted matrix $(M, \ell) \in Y$. Furthermore, there exists $(M', \ell') \in \langle X \rangle$ such that $(M, \ell) \preccurlyeq_{\frac{\varepsilon}{2}} (M', \ell')$. This implies $\frac{1}{\ell} M \leqslant \frac{1}{\ell'} M' + \frac{\varepsilon}{2}$. It follows that $I \otimes M' \otimes T > (1-\varepsilon)\ell'$. This contradicts Lemma 9.
- Assume $f \not\leqslant g$, then by Lemma 9, there exists a matrix $M \in \langle X \rangle$ such that $I \otimes \frac{1}{\ell} M \otimes T > 1$. Furthermore, there exists $M' \in Y$ such that $(M, \ell) \preceq_{\frac{\varepsilon}{2}} (M', \ell')$. This implies $\frac{1}{\ell} M \leqslant \frac{1}{\ell'} M' + \frac{\varepsilon}{2}$, and hence $I \otimes \frac{1}{\ell'} M' \otimes T > 1 - \frac{\varepsilon}{2}$. The algorithm answers "no". $\square$

5 Conclusion and further remarks

In this paper, we have provided an algorithm for deciding the approximate comparison of distance automata. This algorithm involves the computation of the closure under product of sets of–what we call–weighted matrices. This result can be of independent interest.

The main open question is the complexity of the problem. It is clear that the problem is at least PSPACE hard. A correct implementation of the arguments in this paper show that EXSPACE is an upper bound. We do not know what is the exact complexity.

References

1. Mikolaj Bojańczyk and Thomas Colcombet. Bounds in ω-regularity. In *LICS 06*, pages 285–296, 2006.
2. Thomas Colcombet. The theory of stabilisation monoids and regular cost functions. In *Automata, languages and programming. Part II*, volume 5556 of *Lecture Notes in Comput. Sci.*, pages 139–150. Springer, Berlin, 2009.
3. Thomas Colcombet. Green's relations and their use in automata theory. In Adrian Horia Dediu, Shunsuke Inenaga, and Carlos Martín-Vide, editors, *LATA*, volume 6638 of *Lecture Notes in Computer Science*, pages 1–21. Springer, 2011. Invited lecture.
4. Kosaburo Hashiguchi. Limitedness theorem on finite automata with distance functions. *J. Comput. Syst. Sci.*, 24(2):233–244, 1982.
5. Daniel Kirsten. Distance desert automata and the star height problem. *RAIRO*, 3(39):455–509, 2005.

6. Daniel Krob. The equality problem for rational series with multiplicities in the tropical semiring is undecidable. *Internat. J. Algebra Comput.*, 4(3):405–425, 1994.
7. Manfred Kufleitner. A proof of the factorization forest theorem. Technical Report Nr. 2007/05, Universität Stuttgart, Germany, 2007.
8. Hing Leung. On the topological structure of a finitely generated semigroup of matrices. *Semigroup Forum*, 37:273–287, 1988.
9. Marcel-Paul Schützenberger. On the definition of a family of automata. *Information and Control*, 4:245–270, 1961.
10. Imre Simon. Piecewise testable events. In H. Brackage, editor, *Proc. 2nd GI Conf.*, volume 33, pages 214–222. Springer, 1975.
11. Imre Simon. On semigroups of matrices over the tropical semiring. *RAIRO ITA*, 28(3-4):277–294, 1994.

In Section A, we give the complete proof of Lemma 2 and in Section B the one of Lemma 3. These two lemmas are stated in Section 3 to prove Theorem 5. Section C is devoted to give the proof of Lemma 8, stated in Section 4.

A Proof of Lemma 2

Actually, in this part, we prove a generalisation of Lemma 2. We consider a product of p weighted matrices, for an integer p, (and not only a product of 2 weighted matrices). Lemma 12 exhibits a finitely presented set that approximates a product of p finitely presented sets. First, we state two lemmas used in the proof. We state them in the general case too. (Lemma 2 only needs the case where $p = 2$, but the proof of Lemma 3 uses Lemmas 10, 11 and 12 in a more general way.)

First, Lemma 10 says that given a bounded number of weighted matrices, it is possible to bound the difference between two products in these matrices, namely between $(\ell_1 M_1 \otimes \cdots \otimes \ell_p M_p, \ell)$ and $(\ell'_1 M_1 \otimes \ell'_2 M_2 \otimes \cdots \otimes \ell'_p M_p, \ell)$, provided that coefficients ℓ_i and ℓ'_i were sufficiently close.

Lemma 10. *For all $\varepsilon > 0$, for all $a \geqslant 0$, for all positive integer p, there is $\eta > 0$ such that for all integers $\ell_1, \ell_2, \ldots, \ell_p$, $\ell'_1, \ell'_2, \ldots, \ell'_p$, ℓ, and for all matrices $M_1, M_2, \ldots, M_p$ whose greatest entry is a, we have:*

$$\textit{if for all } i, |\ell_i - \ell'_i| \leqslant \eta\ell \textit{ then:}$$

$$(\ell_1 M_1 \otimes \ell_2 M_2 \otimes \cdots \otimes \ell_p M_p, \ell) \preccurlyeq_\varepsilon (\ell'_1 M_1 \otimes \ell'_2 M_2 \otimes \cdots \otimes \ell'_p M_p, \ell)\ .$$

Proof. Set $\eta = \varepsilon/pa$, we have:

$$\begin{aligned}
\ell_1 M_1 \otimes \cdots \otimes \ell_p M_p &\leqslant (\ell'_1 + \eta\ell) M_1 \otimes \cdots \otimes (\ell'_p + \eta\ell) M_p \\
&\leqslant (\ell'_1 M_1 + a\eta\ell) \otimes \cdots \otimes (\ell'_p M_p + a\eta\ell) \\
&\leqslant \ell'_1 M_1 \otimes \cdots \otimes \ell'_p M_p + (pa\eta)\ell.
\end{aligned}$$

□

As for Lemma 11, it deals with the products of the sets $\{(\ell M_i, \ell), \ell \geqslant x_i\}$. It states that, provided that ℓ is large enough, we obtain all the products ("barycentre") $(\ell(\lambda_1 M_1 \otimes \cdots \otimes \lambda_p M_p), \ell)$ with $\sum_{i=1}^p \lambda_i = 1$.

Lemma 11. *Let $x_1, x_2, \ldots, x_p$ be positive integers. For all $\eta > 0$, there is a positive integer z such that, for all $k \geqslant z$, for all $0 \leqslant \lambda_1, \lambda_2, \ldots, \lambda_p \leqslant 1$ with $\sum_{i=1}^p \lambda_i = 1$, for all i, there are integers $y_i \geqslant x_i$ such that $\sum_{j=1}^p y_j = k$ and $|y_i - \lambda_i k| \leqslant \eta k$.*

Proof. Let $x_1, x_2, \ldots, x_p$ be positive integers and $\eta > 0$. Denote by x the positive integer $\max_{\{1 \leqslant i \leqslant p\}} x_i$. Set z a positive integer such that $z \geqslant \frac{2px}{\eta}$ and $z > 2p^2 x$. Let $k \geqslant z$ and $0 \leqslant \lambda_1, \lambda_2, \ldots, \lambda_p \leqslant 1$ with $\sum_{i=1}^p \lambda_i = 1$.

- If $\lambda_i \leqslant \frac{x}{k}$, set $y_i = x$. Set Γ the set of such indices.
- If $\frac{x}{k} < \lambda_i \leqslant \frac{2px}{k}$, then there is $a_i > 0$ such that:

$$\lambda_i - \eta \leqslant \frac{a_i x}{k} \leqslant \lambda_i \leqslant \frac{(a_i+1)x}{k} \leqslant \lambda_i + \eta$$

  (since $\frac{x}{k} \leqslant \eta$). Set Γ' the set of such indices.
- If $\lambda_i > \frac{2px}{k}$, then there is $a_i > 0$ such that:

$$\lambda_i - \eta \leqslant \frac{a_i px}{k} \leqslant \frac{(a_i+1)px}{k} \leqslant \lambda_i \leqslant \frac{(a_i+2)px}{k} \leqslant \lambda_i + \eta$$

  (since $\frac{px}{k} \leqslant \frac{\eta}{2}$). Set Γ'' the set of such indices and remark that $|\Gamma''| \geqslant 1$ (since $\sum_{i=1}^{p} \lambda_i = 1$, and $\frac{2p^2x}{k} > 1$).

Since $\sum_{i=1}^{p} \lambda_i = 1$, then

$$\begin{aligned}
\sum_{i\in\Gamma'} a_i x + \sum_{i\in\Gamma''} a_i px &\leqslant \sum_{i\in\Gamma'} a_i x + \sum_{i\in\Gamma''} (a_i+1)px && (1)\\
&\leqslant k \\
&\leqslant |\Gamma|x + \sum_{i\in\Gamma'} (a_i+1)x + \sum_{i\in\Gamma''} (a_i+2)px \ . && (2)
\end{aligned}$$

Now, let $q = k - |\Gamma|x - \sum_{i\in\Gamma'} a_i x - \sum_{i\in\Gamma''} a_i px$, thanks to (2), observe that $q \leqslant |\Gamma'|x + 2|\Gamma''|px$ and since $|\Gamma''| \geqslant 1$ and thanks to (1), also observe that $q \geqslant 0$.

One can write $q = 2px\alpha + \beta$ with $\beta \leqslant 2px$. If $\alpha < |\Gamma''|$, then for $i \in \Gamma'$ set $y_i = a_i$, for α indices in Γ'' set $y_i = (a_i+2)px$, for one index in Γ'' set $y_i = a_i px + \beta$, and for the other ones set $y_i = a_i px$.

If $\alpha \geqslant |\Gamma''|$, then we can rewrite $q = 2px|\Gamma''| + \gamma$ with $\gamma \leqslant |\Gamma'|x$. For $i \in \Gamma''$, set $y_i = (a_i+2)px$, and for $i \in \Gamma'$, if $\gamma = |\Gamma'|x$ then set $y_i = (a_i+1)x$, otherwise one can write $\gamma = \delta x + \mu$ with $\delta < |\Gamma'|$ and $\mu < x$. For δ indices in Γ' set $y_i = (a_i+1)x$, for one index in Γ' set $y_i = a_i x + \mu$, and for the other ones set $y_i = a_i x$.

One can easily check that the integers y_i satisfy the conditions. □

Lemma 12 (generalisation of Lemma 2). *For all $\varepsilon > 0$, and finitely presented sets $X_1, \dots, X_p \subseteq \mathcal{W}_{n,n}$, there is a computable and finitely presented set Z such that:*

$$Z \approx_\varepsilon X_1 \otimes \cdots \otimes X_p \ .$$

Proof. Assume Lemma 12 is true for $p = 2$, then by induction, using Lemma 1, it will be true for all the integers p. Let us prove Lemma 12 for $p = 2$. Let $\varepsilon > 0$ and X and Y be two sets as in the lemma. Since the finitely presented sets of weighted matrices are closed under union, it is sufficient to prove Lemma 12 for the atomic blocks of the finite presentation.

– If $X = \{(M, \ell)\}$ and $Y = \{(M', \ell')\}$, then we can choose $Z = \{(M \otimes M', \ell + \ell')\}$.

– If $X = \{(M, \ell)\}$ and $Y = \{(yN, y) \mid y \geqslant k\}$, set a the greatest coefficient of $\frac{1}{\ell}M$ and N. For $p = 2$, consider η given by Lemma 10. Set z an integer such that $z \geqslant k + \ell$ and $z \geqslant \frac{1-\eta}{\eta}\ell$. Set $Z = Z_1 \cup Z_2$ with

$$Z_1 = \bigcup_{k \leqslant y < z} \{(M \otimes yN, \ell + y)\}$$
$$\text{and } Z_2 = \{(y(\widetilde{M \otimes N}), y) \mid y \geqslant \ell + z\} \ .$$

Then Z is finitely presented. We only need to prove that $X \otimes Y \approx_\varepsilon Z$.

- Let $(M, \ell) \otimes (yN, y) = (M \otimes yN, \ell + y) \in X \otimes Y$ with $y \geqslant k$. If $k \leqslant y < z$, then $(M, \ell) \otimes (yN, y) \in Z_1$. If $y \geqslant z$ then $\ell \leqslant \eta(y + \ell)$, thus by Lemma 10, we get:

$$(M, \ell) \otimes (yN, y) \preccurlyeq_\varepsilon (\widetilde{M} \otimes (y + \ell)N, y + \ell) = ((y + \ell)(\widetilde{M \otimes N}), y + \ell) \in Z_2 \ .$$

- Conversely, first $Z_1 \subseteq X \otimes Y$. Furthermore, for $y \geqslant \ell + z$ we have $(y(\widetilde{M \otimes N}), y) = (\widetilde{M} \otimes yN, y)$ and by Lemma 10,

$$(\widetilde{M} \otimes yN, y) \preccurlyeq_\varepsilon (M \otimes (y - \ell)N, y) = (M, \ell) \otimes ((y - \ell)N, y - \ell) \in X \otimes Y \ .$$

For $X = \{(xM, x) \mid x \geqslant \ell\}$ and $Y = \{(N, k)\}$, it is the same thing.

– If $X = \{(xM, x) \mid x \geqslant \ell\}$ and $Y = \{(yN, y) \mid y \geqslant k\}$, let a be the greatest coefficient of M and N and consider η given by Lemma 10 (for $p = 2$). Set z the integer given by Lemma 11 for $p = 2$, $x_1 = \ell$ and $x_2 = k$ and η. Finally, set $Z = Z_1 \cup Z_2$ with:

$$Z_1 = \{(xM \otimes yN, x + y) \mid \ell \leqslant x < z,\ k \leqslant y < z\}$$
$$\text{and } Z_2 = \bigcup_{\lambda \in ([0,1] \cap \eta\mathbb{N})} \{(t(\lambda M \otimes (1 - \lambda)N), t) \mid t \geqslant z\} \ .$$

Z is finitely presented and $X \otimes Y \approx_\varepsilon Z$. Indeed:

- let $(xM, x) \otimes (yN, y) \in X \otimes Y$. If $x < z$ and $y < z$, $(xM, x) \otimes (yN, y) \in Z_1$. Otherwise, $(xM, x) \otimes (yN, y) = ((x + y)(\frac{x}{x+y}M \otimes \frac{y}{x+y}N), x + y)$ with $x + y \geqslant z$. Then there is $\lambda \in ([0, 1] \cap \eta\mathbb{N})$ such that $|\frac{x}{x+y} - \lambda| \leqslant \eta$ and thus $|\frac{y}{x+y} - (1 - \lambda)| \leqslant \eta$, so by Lemma 10, $(xM, x) \otimes (yN, y) \preccurlyeq_\varepsilon ((x + y)(\lambda M \otimes (1 - \lambda)N), x + y) \in Z_2$.
- First $Z_1 \subseteq X \otimes Y$. Let $(t(\lambda M \otimes (1 - \lambda)N), t) \in Z_2$ with $t \geqslant z$. By Lemma 11, there are $x \geqslant \ell$ and $y \geqslant k$ such that $x + y = t$, $|x - \lambda t| \leqslant \eta t$ and $|y - (1 - \lambda)t| \leqslant \eta t$. By Lemma 10, we get:

$$(t(\lambda M \otimes (1 - \lambda)N), t) \preccurlyeq_\varepsilon ((x + y)(\frac{x}{x + y}M \otimes \frac{y}{x + y}N), x + y) \in X \otimes Y \ .$$

□

B Proof of Lemma 3

In this part, we prove Lemma 3 that exhibits a finitely presented set that approximates the closure of a finitely presented set in which all the matrices have the same idempotent projection ($i.e$ there is an idempotent matrix E such that for all (M, ℓ), $\widetilde{M} = E$).

There are three steps in the proof. In the first step (Lemma 17), we split a product of weighted matrices into "simpler" products (in a sense given in Subsection 3.3). Lemmas 13, 14, 15, 16 are required to prove it. In the second step, we show that some of these "simpler" products are ε-equivalent to finitely presented sets, and other ones are ε-equivalent to the closure under products of finitely presented and uniform sets. This is the aim of Lemma 18. The third step is devoted to give a finitely presented set ε-equivalent to the closure of finitely presented and uniform sets. It is the aim of Lemma 19.

Lemma 13. *The function:*

$$f : (x_1, x_2, \ldots, x_r) \mapsto \sum_{i=1}^{r} x_i(1 - x_{i-1}) \cdots (1 - x_1)$$

is increasing over $[0, 1]^r$ *(in the sense that if for all* i*,* $x_i \leqslant y_i$*, then* $f(x_1, x_2, \ldots, x_r) \leqslant f(y_1, y_2, \ldots, y_r)$*).*

Proof. For all i, we have $\frac{\partial f}{\partial x_i}(x_1, \ldots, x_r) = \prod_{j \neq i}(1 - x_j) \geqslant 0$. □

Lemma 14. *Let* $(b_0, b_1, \ldots, b_p)$ *some positive numbers such that* $\sum_{i=0}^{p} b_i = \ell$*. For* $i \geqslant 1$*, set* $c_i = \frac{b_i}{b_0 + \ldots + b_i}$*, then:*

$$b_p + \ldots + b_1 = \ell f(c_p, c_{p-1}, \ldots, c_1) \ .$$

Proof. We prove by induction that for all i, $b_i = \ell c_i(1 - c_{i+1}) \cdots (1 - c_p)$. The result follows. □

Lemma 15. *Let* $\eta > 0$*, there exists an integer* p *such that, given a positive integer* r *and* $a_1, \ldots, a_r$ *non-negative numbers, there are* $0 = i_0 < i_1 < i_2 < \ldots < i_k < i_{k+1} = r$ *such that:*

- *for all* $j \in \{1, \ldots, k\}$*,* $\frac{a_{i_j}}{a_{i_{j-1}+1} + \ldots + a_{i_j}} \leqslant \eta$*,*
- *for all* $j \in \{1, \ldots, k+1\}$*, there is an integer* $s \leqslant p$ *and a set* Γ *of* s *indices between* $i_{j-1} + 1$ *and* i_j *such that* $\frac{\sum_{i \in \Gamma} a_i}{a_{i_{j-1}+1} + \ldots + a_{i_j}} \geqslant 1 - \eta$*.*

Proof. Let $\eta > 0$, set p an positive integer such that $(1 - \eta)^{p-1} \leqslant \eta$. Define i_j by induction on j. Let $i_0 = 0$ and then, suppose we have computed i_j. We set i_{j+1} the smallest index greater than i_j such that $\frac{a_{i_{j+1}}}{a_{i_j+1} + \ldots + a_{i_{j+1}}} \leqslant \eta$ if it exists, and $i_{j+1} = i_{k+1} = r$ otherwise. If $i_j - i_{j-1} \leqslant p$ then take Γ all the indices between $i_{j-1} + 1$ and i_j. Otherwise, take $\Gamma = \{i_j - p + 1, \ldots, i_j\}$. Set $\ell = \sum_{i=i_{j-1}+1}^{i_j} a_i$

and $b = \sum_{i=i_{j-1}+1}^{i_j-p} a_i$. We are going to prove that: $\frac{\sum_{i\in\Gamma} a_i}{\ell} \geqslant 1-\eta$. By the definition of i_j, we have for all $i_j - p + 1 \leqslant i < i_j$ that $c_i = \frac{a_i}{b+a_{i_j-p+1}+\ldots+a_i} > \eta$. Now apply Lemma 14 to $(b, a_{i_j-p+1}, \ldots, a_{i_j-1})$, we get

$$\sum_{i\in\Gamma} a_i = (\ell - a_{i_j}) f(c_{i_j-1}, \ldots, c_{i_j-p+1}) + a_{i_j} \ .$$

Besides by Lemma 13,

$$f(c_{i_j-1}, \ldots, c_{i_j-p+1}) \geqslant f(\eta, \ldots, \eta) = 1 - (1-\eta)^{p-1}.$$

And hence we get:

$$\sum_{i\in\Gamma} a_i \geqslant \ell(1-\eta) \ .$$

□

Lemma 16. *Let $\eta > 0$, there exists an integer p such that, given a positive integer r and $a_1, \ldots, a_r$ non-negative numbers, there are $0 = i_0 < i_1 < i_2 < \ldots < i_k < i_{k+1} = r$ such that:*

- *for all $j \in \{1, \ldots, k\}$, $\frac{a_{i_j+1}}{a_{i_j+1}+\ldots+a_{i_{j+1}}} \leqslant \eta$,*
- *for all $j \in \{1, \ldots, k\}$, $\frac{a_{i_j}}{a_{i_{j-1}+1}+\ldots+a_{i_j}} \leqslant \eta$,*
- *for all $j \in \{0, \ldots, k\}$, there is $s \leqslant p$ and a set Γ of s indices between i_j+1 and i_{j+1} such that $\frac{\sum_{a\in\Gamma} a}{a_{i_j+1}+\ldots+a_{i_{j+1}}} \geqslant 1-\eta$.*

Proof. Let $\eta > 0$. Let $0 < \gamma \leqslant \eta$ such that $2\gamma - \gamma^2 \leqslant \eta$. Consider the integer given by Lemma 15 for γ, and set p its square. Let r be a positive integer and $a_1, \ldots, a_r$ non-negative numbers, Lemma 15 for γ gives a sequence of indices $0 = q_0 < q_1 < q_2 < \ldots < q_k < q_{k+1} = r$ satisfying some conditions. For all $j \in \{1, \ldots, k+1\}$, set $b_{k+2-j} = a_{q_{j-1}+1} + \ldots + a_{q_j}$. Now apply Lemma 16 for γ on $b_1, b_2, \ldots, b_{k+1}$. It gives another sequence of indices $0 = j_0 < j_1 < j_2 < \ldots < j_m < j_{m+1} = k+1$ satisfying some conditions. For $t \in \{0, \ldots, m+1\}$, set $i_{m+1-t} = q_{k+1-j_t}$. Then indices i_t satisfy the lemma. □

Lemma 17 (Lemma 5 in the abstract). *Let X be a set of weighted matrices, for all $\eta > 0$, there is an integer p such that:*

$$\langle X \rangle = \langle X \rangle_{p,\eta} \otimes \langle \langle X \rangle^u_{p,\eta} \rangle \otimes \langle X \rangle_{p,\eta} \ .$$

Proof. Let $\eta > 0$, consider the integer p given in Lemma 16. We have $\langle X \rangle_{p,\eta} \otimes \langle \langle X \rangle^u_{p,\eta} \rangle \otimes \langle X \rangle_{p,\eta} \subseteq \langle X \rangle$. Conversely, if $(M_1, \ell_1) \otimes \cdots \otimes (M_r, \ell_r) \in \langle X \rangle$, apply Lemma 16 on the sequence of weights $\ell_1, \ldots, \ell_r$. We obtain indices $0 = i_0 < i_1 < i_2 < \ldots < i_k < i_{k+1} = r$ such that for $j = 0, \ldots, k$, the products $(M_{i_j+1}, \ell_{i_j+1}) \otimes \cdots \otimes (M_{i_{j+1}}, \ell_{i_{j+1}})$ have the good form. □

Lemma 18 (Lemmas 4 and 6 in the abstract). *For all $\gamma > 0$, for all $a \geqslant 0$, there exists $\eta > 0$ such that for all finitely presented set X with $\widetilde{X} = \{E\}$ for an idempotent matrix E and such that for all $(M, \ell) \in X$, entries of $\frac{1}{\ell}M$ are smaller than a, and for all p, there exist effectively Y and Z finitely presented such that:*

$$Y \approx_\gamma \langle X \rangle_{p,\eta} \text{ and } Z \approx_\gamma \langle X \rangle^u_{p,\eta} \ .$$

Moreover, all weighted matrices in Z can be uniform.

Proof. Let $\gamma > 0$ and $a \geqslant 0$, set $\eta = \frac{\gamma}{2a}$. If X is finite, let ℓ the greatest weight of a matrix in X, set $Y' = X^{\frac{p(\eta(\ell-1)+1)}{1-\eta}}$. Then $\langle X \rangle_{p,\eta} \subseteq Y'$. (Indeed, if $P \in \langle X \rangle_{p,\eta}$, let $x + p$ the number of matrices in this product (if it is greater or equal to p), $\ell_1, \ldots, \ell_x$ the weights of the matrices that "do not count", and $k_1, \ldots, k_p$ the weights of the matrices that "count". Then, $\frac{\ell_1+\ldots+\ell_x}{\ell_1+\ldots+\ell_x+k_1+\ldots+k_p} \leqslant \eta$,

$$\frac{1}{\eta} \leqslant 1 + \frac{k_1 + \ldots + k_p}{\ell_1 + \ldots + \ell_x} \leqslant 1 + \frac{p\ell}{x}$$

that implies $x + p \leqslant \frac{p(\eta(\ell-1)+1)}{1-\eta}$.) Thus, $\langle X \rangle_{p,\eta}$ is finite (so finitely presented) and computable.

If X is infinite, let ℓ be the smallest weight of a matrix in X and set k an integer such that $k \geqslant \frac{(p+1)\ell}{\eta}$. Let $X_1 = \{(M, t) \in X \mid t \geqslant k\}$, and let $X_2 = \{(M, t) \in X \mid t < k\}$. We have:

$$\langle X \rangle_{p,\eta} = \langle X_2 \rangle_{p,\eta} \cup X'$$

where $X' = \langle X \rangle_{p,\eta} \cap (X_1 \cup (\langle X \rangle \otimes X_1) \cup (X_1 \otimes \langle X \rangle) \cup (\langle X \rangle \otimes X_1 \otimes \langle X \rangle)$ is the set of products in $\langle X \rangle_{p,\eta}$ containing at least one matrix of weight greater or equal to k. Moreover X_2 is finite, thus, by the arguments given above, $\langle X_2 \rangle_{p,\eta}$ is finite.

Let Y the maximal set such that:

$$Y' \subseteq (X \cup (X \otimes \{(E, \ell)\}) \cup (\{(E, \ell)\} \otimes X) \cup (\{(E, \ell)\} \otimes X \otimes \{(E, \ell)\}))^p$$

and every products in Y' contain at least a matrix in X of weight greater or equal to k. We will prove that $Y' \approx_{\frac{\gamma}{2}} X'$, and there is Y finitely presented such that $Y \approx_{\frac{\gamma}{2}} Y'$. Then by Lemma 1, the conclusion will follow.

Consider a product P in X'. By definition, there is a sequence of weighted matrices in P such that the sum of their weights has a ratio smaller than η. Thus, set P' the product P in which these weighted matrices are replaced by (E, ℓ). Then $P' \in Y'$ and $P \preccurlyeq_{\eta a} P'$. Hence $X' \preccurlyeq_{\frac{\gamma}{2}} Y'$.

Consider now a product P' in Y'. By definition, there exists (M, ℓ) in X (recall ℓ is the smallest weight in X). Then set P the product P' in which all the matrices (E, ℓ) are replaced by (M, ℓ). Then since there is a matrix in P' (and also in P) of weight greater than k, the sum of the weights of the added matrices is smaller than $\ell(p + 1) \leqslant \eta k$ by definition. Hence $P \in X'$. Moreover, $P' \preccurlyeq_{\eta a} P$.

Finally, $Y' \approx_{\frac{\gamma}{2}} X'$.

For all i, define the set

$$Z'_i = (X \cup (X \otimes \{(E,\ell)\}) \cup (\{(E,\ell)\} \otimes X) \cup (\{(E,\ell)\} \otimes X \otimes \{(E,\ell)\}))^i \ .$$

We can rewrite the definition of Y':

$$Y' = \bigcup_{i=0}^{p-1} Z'_i \otimes X_1 \otimes Z'_{p-i-1} \ .$$

The sets X_1, X, $X \otimes (E,\ell)$, $(E,\ell) \otimes X$ and $(E,\ell) \otimes X \otimes (E,\ell)$ are finitely presented, and so $X \cup (X \otimes \{(E,\ell)\}) \cup (\{(E,\ell)\} \otimes X) \cup (\{(E,\ell)\} \otimes X \otimes \{(E,\ell)\})$ is also finitely presented. Then by Lemma 12, for all i, there are computable and finitely presented sets Z_i such that $Z_i \approx_{\frac{\gamma}{4}} Z'_i$. Besides there is a finitely presented set Y_i such that $Y_i \approx_{\frac{\gamma}{4}} Z_i \otimes X_1 \otimes Z_{p-i-1}$. Finally, $Y = \bigcup_{i=0}^{p-1} Y_i$ is finitely presented and by Lemma 1, $Y' \approx_{\frac{\gamma}{2}} Y$.

Remark that if a product is uniform (with ratio δ) then it is $2\delta a$-equivalent to the same product in which the first and the last matrices are replaced by E. The weighted matrix obtained by doing such product is uniform. By restricting all the sets constructed above, by considering only the products in which the first and last matrices have a ratio smaller than η, the same construction gives us a set Y that is equivalent to a set Z in which all the matrices are uniform. □

Remark 1. We will explain what is the structure of a set of uniform matrices. Suppose M is uniform and $\widetilde{M} = E$. Then, for two states i and j, we say that i and j are connected if $E_{i,j} = E_{j,i} = 0$. First remark that this relation is transitive since E is idempotent.

The main property we have is that if i and j (resp. i' and j') are connected then $M_{i,i'} = M_{j,j'}$ (since $E \otimes M \otimes E = M$). Then two connected states play exactly the same role. Then if M' is also a uniform matrix, (with $\widetilde{M'} = E$), we have $(M \otimes M')_{i,i} = M_{i,i} + M'_{i,i}$.

Finally, remark that for all positive integer k, $kM \leqslant M^k$.

Lemma 19 (Lemma 7 in the abstract). *For all $\eta > 0$, for all finitely presented set X of uniform matrices such that $\widetilde{X} = \{E\}$ for an idempotent matrix E, there exists effectively Z finitely presented such that:*

$$\langle X \rangle \approx_\eta Z \ .$$

Proof. Let $\eta > 0$, we can write:

$$X = \bigcup_{1 \leqslant i \leqslant p} \{(P_i, p_i)\} \cup \bigcup_{1 \leqslant i \leqslant q} \{(xQ_i, x) \mid x \geqslant q_i\} \ .$$

For $1 \leqslant i \leqslant p$, set $M_i = \frac{1}{p_i} P_i$ and for $1 \leqslant i \leqslant q$, $M_{p+i} = Q_i$. Let a be the greatest coefficient of the matrices M_i and $m = p + q$.

Let $\gamma > 0$ such that $\gamma \leqslant \frac{\eta}{2am^2}$.

Set r a positive integer such that $r \geqslant \frac{2n(a+1)}{\gamma}$.

Lemma 11 gives an integer z', such that for all $k \geqslant z'$, for all $0 \leqslant \lambda_1, \ldots, \lambda_m \leqslant 1$ with $\sum_{i=1}^{m} \lambda_i = 1$, for all $i \leqslant p$, there are integers $y_i \geqslant p_i$ and for all $i \leqslant q$, $y_{p+i} \geqslant q_i$ such that $\sum_{j=1}^{m} y_j = k$ and $|y_i - \lambda_i k| \leqslant \gamma k$.

Let $z \geqslant 2p_1 \cdots p_p r \eta^{-1} z'$.

Set $Z = Z_1 \cup Z_2$ with:

$$Z_1 = \{(N_1, t_1) \otimes \cdots \otimes (N_k, t_k) \mid \forall i, (N_i, t_i) \in X, t_1 + \ldots + t_k < z\}$$

$$\text{and } Z_2 = \bigcup_{\substack{\{(\lambda_1,\ldots,\lambda_m) \mid \\ \lambda_1+\ldots+\lambda_m=1 \\ \lambda_1,\ldots,\lambda_{m-1} \in [0,1] \cap \gamma\mathbb{N}\}}} \{(x\left(\frac{1}{r}\left(\lambda_1 M_1 \otimes \cdots \otimes \lambda_m M_m\right)^r\right), x) \mid x \geqslant z\} \ .$$

The set Z is computable and finitely presented.

First, let us show that $Z \preccurlyeq_\eta \langle X \rangle$. We have $Z_1 \subseteq \langle X \rangle$. Let

$$A = (k\left(\frac{1}{r}\left(\lambda_1 M_1 \otimes \cdots \otimes \lambda_m M_m\right)^r\right), k) \in Z_2 \ .$$

Let k' be an integer such that $p_1 \cdots p_p r k' \leqslant k < p_1 \cdots p_p r(k'+1)$. Since $k \geqslant z$, we have $k' \geqslant z'$. Then, for all $i \leqslant p$, there are integers $y_i \geqslant p_i$ and for all $i \leqslant q$, $y_{p+i} \geqslant q_i$ such that $\sum_{j=1}^{m} y_j = k'$ and $|y_i - \lambda_i k'| \leqslant \gamma k'$.

Then:

$$\begin{aligned}
k\left(\frac{1}{r}\left(\lambda_1 M_1 \otimes \cdots \otimes \lambda_m M_m\right)^r\right) = \quad & \left(\frac{k\lambda_1}{r} M_1 \otimes \cdots \otimes \frac{k\lambda_m}{r} M_m\right)^r \\
\leqslant \quad & (\frac{p_1 \cdots p_p r y_1 + p_1 \cdots p_p r \gamma k' + p_1 \cdots p_p r \lambda_1}{r} M_1 \\
& \otimes \cdots \otimes \\
& \frac{p_1 \cdots p_p r y_m + p_1 \cdots p_p r \gamma k' + p_1 \cdots p_p r \lambda_m}{r} M_m)^r \\
\leqslant \quad & (p_1 \cdots p_p y_1 M_1 \otimes \cdots \otimes p_1 \cdots p_p y_m M_m)^r \\
& + p_1 \cdots p_p a m r \gamma k' + p_1 \cdots p_p r \\
\leqslant \quad & (p_1 \cdots p_p y_1 M_1 \otimes \cdots \otimes p_1 \cdots p_p y_m M_m)^r \\
& + k\frac{\eta}{2} + k\frac{\eta}{2} \\
\leqslant \quad & ((p_1 M_1)^{p_2 \cdots p_p y_1} \otimes \cdots \otimes M_m^{p_1 \cdots p_p y_m})^r + k\eta \ .
\end{aligned}$$

Then $A \preccurlyeq_\eta (((p_1 M_1)^{p_2 \cdots p_p y_1} \otimes \cdots \otimes M_m^{p_1 \cdots p_p y_m})^r, p_1 \cdots p_p k' r) \in \langle X \rangle$.

Now let us show that $\langle X \rangle \preccurlyeq_\eta Z$. Let:

$$A = (\ell_1 M_{i_1}, \ell_1) \otimes \cdots \otimes (\ell_k M_{i_k}, \ell_k) \in \langle X \rangle.$$

Set $\ell = \sum_{i=1}^{k} \ell_i$, and $\lambda_i = \frac{\sum_{i_j = i} \ell_j}{\ell}$. If $\ell < z$ then $A \in Z_1$. Otherwise, there is $(\lambda'_i)_{i=1,\ldots,m}$ such that for all $i = 1, \ldots, m-1$, $\lambda'_i \in \gamma\mathbb{N}$, $\sum_{i=1}^{m} \lambda'_i = 1$ and for all $i = 1, \ldots, m$, $|\lambda_i - \lambda'_i| \leqslant m\gamma$.

Set $B = (\ell\left(\frac{1}{r}\left(\lambda'_1 M_1 \otimes \cdots \otimes \lambda'_m M_m\right)^r\right), \ell)$. Then $B \in Z_2$. Let us prove that $A \preccurlyeq_\eta B$.

First, $\ell\left(\frac{1}{r}\left(\lambda_1 M_1 \otimes \cdots \otimes \lambda_m M_m\right)^r\right) \leqslant \ell\left(\frac{1}{r}\left(\lambda'_1 M_1 \otimes \cdots \otimes \lambda'_m M_m\right)^r\right) + m^2\gamma a\ell$, thus $\ell\left(\frac{1}{r}\left(\lambda_1 M_1 \otimes \cdots \otimes \lambda_m M_m\right)^r\right) \leqslant \ell\left(\frac{1}{r}\left(\lambda'_1 M_1 \otimes \cdots \otimes \lambda'_m M_m\right)^r\right) + \frac{\eta}{2}\ell$. Then we only need to prove that:

$$\ell_1 M_{i_1} \otimes \cdots \otimes \ell_k M_{i_k} \leqslant \ell\left(\frac{1}{r}\left(\lambda_1 M_1 \otimes \cdots \otimes \lambda_m M_m\right)^r\right) + \frac{\eta}{2}\ell \ .$$

Let α and β be two states. Let ξ be one of the states such that $E_{\alpha,\xi} = E_{\xi,\beta} = 0$, and $(\lambda_1 M_1 \otimes \cdots \otimes \lambda_m M_m)_{\xi,\xi}$ is minimal. Let us call μ this quantity. Thanks to Remark 1, $\mu = (\lambda_1 M_1)_{\xi,\xi} + \ldots + (\lambda_m M_m)_{\xi,\xi}$.

A product of uniform matrices is uniform, ($E \otimes M \otimes M' \otimes E = E \otimes E \otimes M \otimes E \otimes E \otimes M' \otimes E \otimes E = M \otimes M'$) then:

$$\begin{aligned}(\ell_1 M_{i_1} \otimes \cdots \otimes \ell_k M_{i_k})_{\alpha,\beta} &= (E \otimes \ell_1 M_{i_1} \otimes \cdots \otimes \ell_k M_{i_k} \otimes E)_{\alpha,\beta} \\ &\leqslant E_{\alpha,\xi} + (\ell_1 M_{i_1})_{\xi,\xi} + \ldots + (\ell_k M_{i_k})_{\xi,\xi} + E_{\xi,\beta} \\ &\leqslant \ell\mu \ .\end{aligned}$$

Finally, we show that $\ell\mu \leqslant \ell\left(\frac{1}{r}\left(\lambda_1 M_1 \otimes \cdots \otimes \lambda_m M_m\right)^r\right) + \frac{\eta}{2}\ell$. For s large enough (namely greater than $\frac{n(a+1)}{\gamma}$), $(s-n)\mu \leqslant (\lambda_1 M_1 \otimes \cdots \otimes \lambda_m M_m)^s_{\alpha,\beta} \leqslant s\mu$. (It is due to the structure of a set of uniform matrices.) Then, $\ell\mu - \ell n\mu r^{-1} \leqslant \ell\left(\frac{1}{r}\left(\lambda_1 M_1 \otimes \cdots \otimes \lambda_m M_m\right)^r\right)$, that is why $\ell\mu \leqslant \ell\left(\frac{1}{r}\left(\lambda_1 M_1 \otimes \cdots \otimes \lambda_m M_m\right)^r\right) + \frac{\eta}{2}\ell$. ☐

Lemma 20 (Lemma 3 in the abstract). *For all $\varepsilon > 0$, for all finitely presented set X such that $\widetilde{X} = \{E\}$ for an idempotent matrix E, there exists effectively Z finitely presented such that:*

$$Z \approx_\varepsilon \langle X \rangle \ .$$

Proof. Let $\varepsilon > 0$, let η given by Lemma 18 applied to $\frac{\varepsilon}{4}$.

By Lemma 17, there is p such that:

$$\langle X \rangle = \langle X \rangle_{p,\eta} \otimes \langle \langle X \rangle^u_{p,\eta} \rangle \otimes \langle X \rangle_{p,\eta} \ .$$

By Lemma 18, there is T and V finitely presented such that $\langle X \rangle_{p,\eta} \approx_{\frac{\varepsilon}{4}} T$ and $\langle X \rangle^u_{p,\eta} \approx_{\frac{\varepsilon}{4}} V$. Moreover, all the weighted matrices in V are uniform. Then by Lemma 1,

$$\langle X \rangle \approx_{\frac{\varepsilon}{4}} T \otimes \langle V \rangle \otimes T \ .$$

By Lemma 19, there is a finitely presented set Y, such that $\langle V \rangle \approx_{\frac{\varepsilon}{4}} Y$. Then by Lemma 1,

$$\langle X \rangle \approx_{\frac{\varepsilon}{2}} T \otimes Y \otimes T \ .$$

Finally, by Lemma 12, there is Z finitely presented such that $T \otimes Y \otimes T \approx_{\frac{\varepsilon}{2}} Z$.

We conclude by Lemma 1 that $\langle X \rangle \approx_\varepsilon Z$. ☐

C The reduction

Lemma 21 (Lemma 8 in the abstract). *There are* $(n+2, n+2)$*-matrices* $(C^{p,q})_{p,q\in G}$ *over* $\overline{\mathbb{R}^+}$ *and vectors* I_g *and* T_g *such that for all* $p_1, q_1, \dots, p_k, q_k \in Q_g$*,*

$$I_C \otimes C^{p_1,q_1} \otimes \cdots \otimes C^{p_k,q_k} \otimes T_C$$
$$= \begin{cases} \infty & \text{if } p_0 \in I_g,\ q_0 = p_1,\ \dots,\ q_{k-1} = p_k \text{ and } q_k \in T_g, \\ 0 & \text{otherwise.} \end{cases}$$

This is implemented in matrix form as follows. For each p, q where $p, q \in Q_g$, set the matrix $C^{p,q}$ that has indices in $Q_g \cup \{i, \perp\}$, to be such that:

$$(C^{p,q})_{p',q'} = \begin{cases} 0 & \text{if } p' = i,\ p \in I_g \text{ and } q' = q, \\ 0 & \text{if } p' = i,\ p \notin I_g \text{ and } q' = \perp, \\ 0 & \text{if } p' = p \text{ and } q' = q, \\ 0 & \text{if } p' \neq i \text{ and } p' \neq p \text{ and } q' = \perp, \\ \infty & \text{otherwise.} \end{cases}$$

Define furthermore I_C be the vector with all entries ∞ but i which if 0, and let T_C be the vector with all entries equal to ∞ but T_g (and i if there is an initial and finite state in g).